

NEUVIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

ALGORITHMES ADAPTATIFS POUR LE CODAGE
DU SIGNAL TELEPHONIQUE A 32 kbit/s

Maurice BELLANGER, Jean-Marie RAULIN et Cengiz EVCI

T.R.T. 5, Av. Réaumur - 92350 LE PLESSIS-ROBINSON

RESUME

Le codage MIC Différentiel Adaptatif, appliqué au signal téléphonique, permet d'atteindre avec le débit de 32 kbit/s, un niveau de qualité comparable à celui du codage MIC à 64 kbit/s. Il en résulte la possibilité de doubler la capacité de certaines artères des réseaux de télécommunications.

Deux types d'algorithmes adaptatifs pour le filtre de prédiction dans un tel codage ont été proposés en vue d'une normalisation internationale. L'un utilise une structure de filtre à réponse impulsionnelle finie décomposé en une cascade de cellules de second ordre. L'autre fait appel à une structure mixte RIF-RII. Après une brève présentation, une comparaison est faite entre ces deux approches, en mentionnant plusieurs points importants qui nécessiteraient des études complémentaires.

SUMMARY

The Adaptive Differential PCM Coding technique, when applied to telephone signals, leads, with a bit rate of 32 kbit/s, to a level of performance similar to that achieved by PCM at the rate of 64 kbit/s. Hence the possibility arises of doubling the capacity of some digital links in communication networks.

Two adaptive algorithms have been proposed for the predictor in such a coder, in view of an international standardization. One uses a finite impulse response filter implemented as a cascade of second order sections. The other one combines FIR and IIR filters. After a short introduction, a comparison is carried out between these two approaches, and several important open questions are pointed out.



I - INTRODUCTION

Dans les réseaux de télécommunications la numérisation du signal téléphonique est faite au débit de 64 kbit/s par voie, conformément à la normalisation internationale élaborée par le Comité Consultatif International Télégraphique et Téléphonique (CCITT) depuis plus de dix ans. La transmission d'un tel débit sur les câbles, faisceaux hertziens et satellites est généralement plus coûteuse que la transmission de voies analogiques, ce qui a freiné l'extension des techniques numériques dans les réseaux.

Les progrès récents dans les techniques de traitement du signal ont permis de montrer que le niveau de qualité exigé par le CCITT peut être atteint dans des conditions économiques avantageuses avec un débit de 32 kbit/s, en faisant appel à la fois à la quantification adaptative et à la prédiction linéaire adaptative. Un tel débit peut conduire à doubler la capacité des artères de transmission, ce qui est particulièrement intéressant pour les plus coûteuses d'entre elles, par exemple les liaisons par satellites ; c'est pourquoi le CCITT étudie activement la possibilité d'une normalisation du codage à 32 kbit/s.

Les caractéristiques de qualité exigées du codage à 64 kbit/s sont décrites dans la référence [1]. L'avis G 712 fixe en particulier des limites aux distorsions d'affaiblissement, de temps de propagation, de linéarité et de quantification, qui peuvent être reprises pour le codage à 32 kbit/s. La valeur du temps de propagation doit également être minimisée, pour tenir compte de l'éventualité de codages et décodages multiples dans une communication. Cette éventualité amène une autre contrainte importante, la non accumulation des dégradations dans les transcodages successifs entre les débits de 32 et 64 kbit/s. En effet, un avantage fondamental de la transmission numérique sur la transmission analogique provient du fait que la qualité des communications devient indépendante de la distance, en raison de la régénération. L'introduction d'une nouvelle technique de codage ne doit pas réduire cet avantage.

La résistance aux erreurs de transmissions est aussi un objectif important. Bien que le taux d'erreur sur les liaisons numériques reste généralement inférieur à 10^{-7} , il peut dans certains cas, les communications par satellites en particulier, atteindre des valeurs nettement plus élevées. De ce point de vue le codage à 32 kbit/s doit être comparable au codage à 64 kbit/s.

Les objectifs ainsi définis conduisent à retenir la technique de Modulation par Impulsions et Codage différentielle et adaptative (MIC-DA) pour le codage du signal téléphonique à 32 kbit/s.

II - LE CODAGE MIC-DA

Le schéma d'un codeur MIC-DA est donné par la figure 1. Le signal à coder est représenté par la suite $x(n)$; à chaque valeur de l'indice n , on soustrait de l'élément $x(n)$ une valeur prédite $\tilde{x}(n)$, et la différence $e(n)$ obtenue, appelée erreur de prédiction, est quantifiée à 4 bits pour fournir la suite $d(n)$ transmises, qui correspond au débit de 32 kbit/s.

Le filtre de prédiction qui fournit la suite $\tilde{x}(n)$, est alimenté par le signal $e_q(n)$, qui résulte de l'opération inverse de la quantification faite sur le signal transmis $d(n)$. En l'absence d'erreurs de transmission, le décodeur MIC-DA

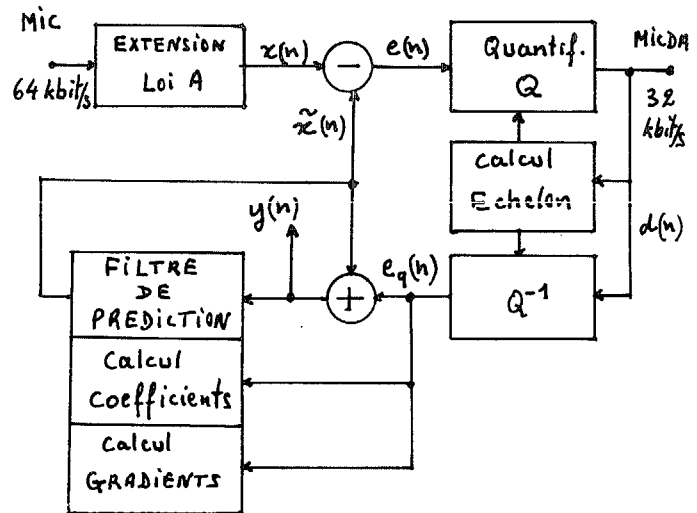


Fig 1 : Codeur MIC Différentiel Adaptatif

fournit la suite $y(n) = \tilde{x}(n) + e_q(n)$, qui en principe ne diffère du signal d'entrée $x(n)$ que par la distorsion de quantification introduite au codeur.

La caractéristique de quantification doit être optimisée de manière à minimiser la distorsion de quantification. Si $e(n)$ est un signal stationnaire dont la distribution des amplitudes suit une loi de Laplace, on montre que cet optimum conduit à un rapport signal à bruit de 18,1 dB [2] ; pour une distribution gaussienne on obtient 20,2 dB. Cependant le signal transmis le plus fréquemment par une voie téléphonique est la parole, dont les propriétés statistiques varient dans le temps et ne peuvent être considérées comme stationnaires que sur des durées de l'ordre de 10 ms.

Il en résulte la nécessité de rendre la quantification adaptative, en estimant en permanence la puissance du signal à quantifier et en faisant varier l'échelon de quantification en conséquence.

Cette estimation doit être faite à partir du signal transmis $d(n)$, pour permettre un fonctionnement correct du décodeur et de manière récursive pour simplifier la réalisation. De plus les variations de l'échelon de quantification $Q(n)$ doivent être exponentielles pour tenir compte de la grande dynamique du signal de parole et rendre la constante de temps d'adaptation relativement indépendante de son amplitude.

Ainsi l'équation d'adaptation de l'échelon de quantification généralement retenue est la suivante :

$$Q(n) = Q(n-1) \cdot M[d(n)] \quad (1)$$

où $M(i)$, avec i entier et $-7 \leq i \leq 8$, est une fonction qui commande la rapidité et la précision de l'adaptation et telle que : $M(i) = M(-i)$; $M(0) < 1$; $M(0) \leq M(1) \leq M(8)$; $M(8) > 1$.

La résistance aux erreurs de transmission est obtenue en introduisant un exposant inférieur à l'unité qui provoque une perte [3]. Finalement l'équation d'adaptation utilisée dans les codeurs MIC-DA a la forme suivante :

$$Q(n) = [Q(n-1)]^{1-\epsilon} M[d(n)] \quad (2)$$

où ϵ est une constante positive de valeur faible, par exemple $\epsilon = \frac{1}{32}$.

La distorsion de quantification introduite étant ainsi liée à la puissance de l'erreur de prédiction à quantifier $e(n)$, il importe de minimiser cette puissance, ce qui peut être obtenu en rendant adaptatif le filtre de prédiction. Cependant une difficulté se présente, car le codeur et le décodeur réalisent des fonctions inverses, comme le montre la figure 2.

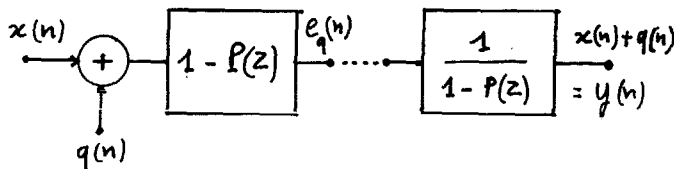


Fig 2:Modèle d'un ensemble Codeur-Décodeur MIC-DA

Dans le codeur, la quantification correspond à l'addition au signal d'entrée $x(n)$ d'un signal d'erreur $q(n)$; la suite $y(n)$ ainsi obtenue est appliquée à un filtre de prédiction, de fonction de transfert $H(Z) = 1 - P(Z)$, qui fournit le signal $e_q(n)$ transmis.

Au décodeur, il faut restituer la suite $y(n)$ par filtrage inverse. Il en résulte la nécessité de contrôler la stabilité, soit au codeur, soit au décodeur, suivant le type de filtre de prédiction, ce qui conduit à des choix de structures appropriées.

Une première structure qui permet un contrôle aisé de stabilité est le treillis. Cependant elle conduit à un nombre important de multiplications et une procédure relativement complexe pour la mise à jour des coefficients ; elle n'a pas été retenue pour le codage du signal téléphonique à 32 kbit/s.

Le filtre à réponse impulsionnelle finie (RIF) en structure cascade au codeur permet également un contrôle aisé de stabilité au décodeur [4].

III-PREDICTION LINEAIRE PAR FILTRE RIF EN CASCADE

La fonction de transfert en Z d'un prédicteur d'ordre P de type RIF s'écrit :

$$H(z) = 1 - \sum_{i=1}^P b_i z^{-i} \quad (3)$$

Conformément à la figure 1, pour minimiser la puissance de la distorsion de quantification, la réalisation correspond à l'expression équivalente suivante :

$$H(z) = \frac{1}{1 + \frac{\sum_{i=1}^P b_i z^{-i}}{1 - \sum_{i=1}^P b_i z^{-i}}} \quad (4)$$

Si le signal $e_q(n)$ est assimilable à un bruit blanc, le gain de prédiction G s'écrit :

$$G^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{dw}{|1 - \sum_{i=1}^P b_i e^{-j\omega} z^{-i}|^2} \quad (5)$$

C'est approximativement le rapport de la puissance du signal d'entrée $x(n)$ à celle de l'erreur de prédiction $e(n)$

Si P est pair, la fonction $H(Z)$ peut se décomposer un produit de $\frac{P}{2}$ facteurs $H_i(Z)$ du second degré et à coefficients réels :

$$H(z) = \prod_{i=1}^{\frac{P}{2}} H_i(z) = \prod_{i=1}^{\frac{P}{2}} (1 - b_{1i} z^{-1} - b_{2i} z^{-2}) \quad (6)$$

Dans ces conditions, on montre que le filtre de prédiction peut être réalisé sous forme cascade, conformément à l'expression suivante [5] :

$$H(z) = \frac{1}{1 + \sum_{j=1}^{\frac{P}{2}} \frac{b_{1j} z^{-1} + b_{2j} z^{-2}}{\prod_{i=1}^j (1 - b_{1i} z^{-1} - b_{2i} z^{-2})}} \quad (7)$$

Le schéma correspondant est donné sur la figure 3, où le signal $e_q(n)$ est appliqué à une cascade de cellules du second ordre purement récursives. Si les pôles de chacune de ces cellules sont complexes,

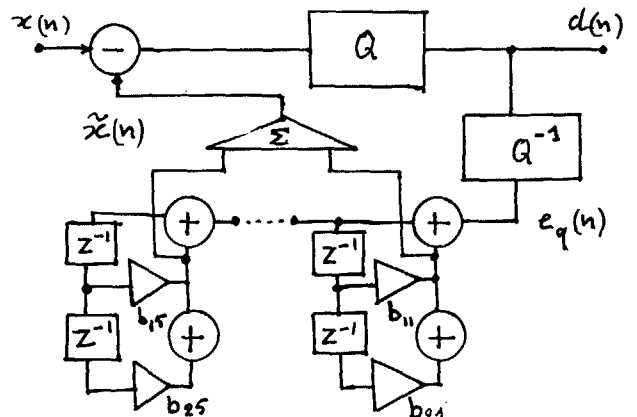


Fig 3:Filtre de prédiction RIF en cascade

la stabilité au décodeur est assurée par la condition:

$$|b_{2i}| < 1 ; \quad 1 \leq i \leq \frac{P}{2}$$



Cette approche permet de bénéficier des avantages connus des cascades de cellules du second ordre, notamment la modularité et la faible précision nécessaire pour les coefficients.

La structure peut être rendue adaptative en utilisant un algorithme du gradient simplifié et en calculant les variations des coefficients par les expressions :

$$d b_{ij}(n) = -\delta \cdot \text{signe} [e(n) \cdot g_{ij}(n)] \quad (8)$$

$$g_{ij}(n) = \frac{\partial e(n)}{\partial b_{ij}} ; j=1,2; i=1, \dots, \frac{P}{2}$$

La constante δ commande la rapidité d'adaptation. Le terme $g_{ij}(n)$ est le gradient correspondant au coefficient $b_{ij}(n)$. On montre que la transformée en Z, de la suite $g_{ij}(n)$, $G_{ij}(Z)$, s'écrit :

$$G_{ij}(Z) = [-z^j / (1 - b_{1i}z^{-1} - b_{2i}z^{-2})] \cdot H(z) \cdot X(z) \quad (9)$$

Ainsi le calcul des gradients met en oeuvre les mêmes cellules récursives que le filtre lui-même, auxquelles est appliquée la suite $e_q(n)$, comme indiqué sur la figure 3.

Les filtres adaptatifs en structure cascade présentent un grand intérêt en raison de leur efficacité et de leur simplicité ; cependant leur mise en oeuvre demande certaines précautions et leurs propriétés et caractéristiques n'ont pas encore été complètement étudiées [6].

Pour le codage du signal téléphonique à 32 kbit/s l'ordre $P = 10$ a été retenu pour le filtre de prédiction, avec la simplification suivante des cellules élémentaires :

$$H_i(z) = 1 + a_i (i z^{-1} + z^{-2}) ; i = -2, -1, 0, 1, 2 \quad (10)$$

Le domaine de variation des coefficients a_i est limité par : $0 \leq a_i < 1$

ce qui assure la stabilité au décodeur.

Les trajectoires des zéros du filtre dans le cercle unité du plan des Z sont données sur la figure 4.

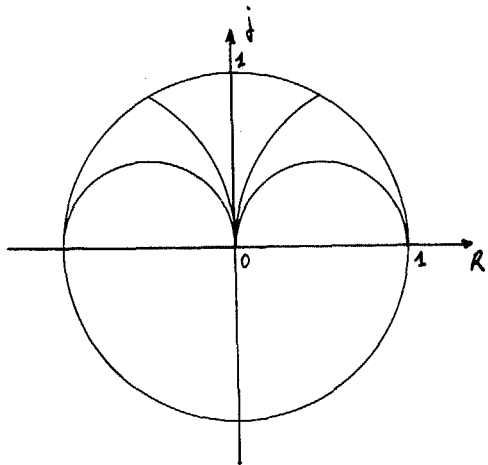


Fig 4 : Zéros du filtre de prédiction

Les zéros se trouvent ainsi séparés et répartis dans la bande des fréquences utiles.

Un équipement utilisant la technique décrite ci-dessus a été réalisé et expérimenté. Il satisfait aux spécifications de l'avis G 712 du CCITT. Pour la parole, le rapport signal à bruit subjectif atteint 34 dB et la résistance aux erreurs de transmission est équivalente à celle du codage MIC à 64 kbit/s ; la propriété de non accumulation des dégradations dans les transcodages successifs est obtenue pour ce type de signal, moyennant certaines précautions dans les opérations d'arrondis.

Avec un signal de modem, à 9600 bit/s, un taux d'erreur de $5 \cdot 10^{-6}$ a été obtenu.

IV - PREDICTION LINEAIRE PAR FILTRAGE RII-RIF

Un filtre de prédiction de type RIF conduit au décodage à un filtre inverse à Réponse Impulsionnelle Infinie (RII). Les transmissions réelles se faisant avec un certain taux d'erreur, le signal ainsi introduit entre le codeur et le décodeur perturbe la transmission, d'autant plus que la réponse impulsionnelle du filtre inverse est plus longue. Pour obtenir une bonne résistance aux erreurs de transmission, il est nécessaire d'introduire un facteur de perte dans la procédure d'adaptation des coefficients ; pour améliorer encore les résultats, on peut envisager de permuter les types de filtre et utiliser une structure RII pour le filtre de prédiction [7-8]. Le gain de prédiction qui peut être obtenu avec un filtre RII est assez limité et en fait il est nécessaire, pour traiter les signaux téléphoniques et en particulier les sonosoides, de conserver une partie non récursive. On aboutit ainsi à une structure mixte.

La fonction de transfert en Z d'un filtre de prédiction mixte RII-RIF s'écrit :

$$H(z) = \frac{1 - \sum_{i=1}^{P_1} b_i z^{-i}}{1 - \sum_{i=1}^{P_2} a_i z^{-i}} \quad (11)$$

Le schéma correspondant est donné par la figure 5

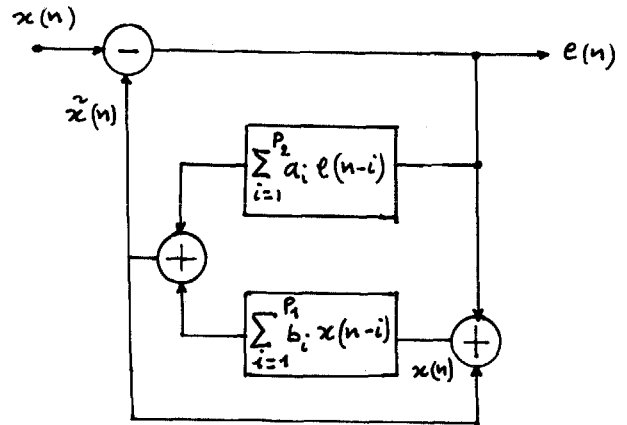


Fig 5 : Filtre de Prédiction mixte pour Codeur MIC-DA.

Si l'ordre P_1 est limité à la valeur 2, le contrôle de stabilité au décodeur est celui d'une cellule du second ordre récurive ; de plus le calcul des gradients correspondant aux coefficients b_1 et b_2 n'est pas nécessaire car ces gradients sont simplement les termes $x(n-1)$ et $x(n-2)$, disponibles dans le filtre, comme l'indique la figure 5.

Dans le filtre RII, la valeur $P_2 = 6$ est suffisante pour les signaux téléphoniques, de parole et données, en complément de la cellule du second ordre de type RIF. L'adaptation des coefficients peut également être faite sans calcul de gradients, en utilisant l'expression suivante :

$$da_i(n) = \delta e(n) \cdot e(n-i); 1 \leq i \leq P_2 \quad (12)$$

Cette partie du filtre ne présente pas de risque d'instabilité au décodage, où elle est de type RIF. Par contre, bien qu'elles n'aient pas été constatées en pratique, des instabilités pourraient théoriquement apparaître au codeur. Ce point devrait faire l'objet d'études complémentaires, de même que les caractéristiques d'adaptation suivant la relation (12).

Dans un codeur MIC-DA à 32 kbit/s le filtre de prédiction mixte ci-dessus conduit à des gains de prédiction très proches des résultats obtenus avec le filtre RIF décrit au paragraphe précédent, pour les signaux de modems et la parole ; l'expérimentation montre qu'il apporte une légère amélioration de la résistance aux erreurs de transmission, conduisant à un codeur MIC-DA qui, pour un signal de modem à 4800 bit/s par exemple, s'avère être, de ce point de vue, meilleur que le codage MIC à 64 kbit/s.

V - CONCLUSION

Les deux types de prédicteurs décrits ont fait l'objet de réalisations dans un codeur MIC Différentiel Adaptatif et sont présentés au CCITT pour une normalisation.

Ces deux approches permettent d'atteindre des caractéristiques de qualité comparables à celles du codage à 64 kbit/s et de satisfaire aux recommandations correspondantes.

Le filtre de prédiction RIF conduit à une structure simple et modulaire, et en particulier il permet de limiter les coefficients à 4 bits ce qui peut conduire à une réalisation sans multiplieurs. Le filtre mixte apporte une légère amélioration de la résistance aux erreurs de transmission, mais nécessite des multiplications de haute précision.

Des résultats expérimentaux importants ont été obtenus. Des compléments d'étude théoriques seraient cependant utiles, sur les filtres adaptatifs en structure cascade et sur les caractéristiques de stabilité et les paramètres d'adaptation des filtres récurifs utilisés comme prédicteurs en codage MIC-DA.

BIBLIOGRAPHIE

- [1] CCITT, Livre jaune, Tome III-3, systèmes de transmission et équipements de multiplexage, Genève, 1981.
- [2] M.D. PAEZ et T.H. GLISSON "Minimum Mean-Squared Error Quantization in Speech PCM and DPCM Systems", IEEE. Trans., COM-20 pp. 225-230, April 1972.
- [3] D.J. GOODMAN and R.M. WILKINSON, "A Robust Adaptive Quantizer", IEEE Trans., COM-23, N°11, pp. 1362-1365, Nov.1975
- [4] L.B. JACKSON and S.L. WOOD, "Linear Prediction in Cascade Form", IEEE Trans., Vol. ASSP-26, Dec. 1978.
- [5] J.M. RAULIN; G. BONNEROT, J.L. JEANDOT, R. LACROIX, "A 60 Channel P M-ADPCM Converter", IEEE Trans., Vol. COM 30, N°4, April 1982.
- [6] P.C. CHING and C.C. GOODYEAR, "Adaptive cascade filter for Speech Analysis", IEE Proc., Vol.130, N°1, Janv. 1983.
- [7] T. NISHITANI, S. AIKOH; T. ARASEKI and . MARUTA, "A 32 kbit/s Toll Quality ADP M Codec using a single chip Signal Processor", Proceedings IASSP 82, Paris 3-5 Mai 1982- pp. 960-963
- [8] D. COINTOT, "A 32 kbit/s ADP M Coder robust to Channel errors", ICASSP 82, Paris 3-5 Mai 1982 pp. 964-967.

